

Ab-Initio Calculations of the Energetics and Kinetics of Defects and Impurities in Semiconductors

Wolfgang Windl

Department of Materials Science and Engineering, The Ohio State University,
Columbus, Ohio 43210-1178, USA

In this paper, we discuss concepts and examples of ab-initio calculations to understand the energetics and kinetics of defects and impurities in silicon. We focus on how to determine equilibrium distributions and diffusion and reaction constants, where modern methods such as the nudged elastic band method in combination with density-functional theory allow a systematic and reliable search for the minimum-energy migration paths and barriers. Since calculations especially for Si self-diffusion give sometimes controversial results when compared to experiments, we also discuss their validity and possible improved theoretical approaches in order to enhance the predictive power of ab-initio calculations.

Introduction

The silicon-based metal oxide semiconductor field effect transistor (MOSFET) is at the heart of today's semiconductor industry. Since the switching speed of MOSFETs increases linearly with shrinking dimensions, which, in turn, also allows for the packing of more MOSFETs on a chip of a given area, the semiconductor industry has managed to constantly improve the performance of computers by continuously scaling a more or less unchanged device geometry. This improvement follows closely Moore's law, which predicts performance doubling approximately every 18 months. Extensive, time consuming and costly engineering work is currently undertaken at semiconductor companies and in academia to further push the miniaturization process of semiconductor devices towards its physical limits and to ultimately define revolutionary new paradigms. In this situation, modeling and simulation has become an increasingly critical cost and especially time saving component of integrated-circuit technology development, provided it is accurate enough.

The shrinking nanoscale dimensions of electronic devices have resulted in an increasing importance of single atoms and their random distributions as well as quantum effects visible in the device performance. Many of these effects cannot be understood anymore within the traditional picture, which is based on classical continuum theory. Process simulation is therefore increasingly being performed by "multiscale modeling", using a hierarchy of tools. On the most fundamental level, the emergence of powerful *ab initio* methods allows us now to gain insight into the physics of mobility and reactions of atoms in the silicon lattice. In the following, we will discuss the most widely used *ab-initio* based techniques and examples of their application to process modeling problems for silicon-based technologies.

Kinetic Parameters from Atomistic Simulations

Whereas the calculation of formation energies – especially for the dilute concentrations in semiconductor devices – is well known and straightforward apart from some corrections that deal with the band-gap problem in the commonly used density-functional approach (1), systematic techniques to determine kinetic parameters have only come into more common use within the past decade.

Diffusion hops or breaking or forming of bonds (reactions) are generally infrequent events. The dynamical evolution of an infrequent-event system consists of vibrational excursions around a (local) minimum-energy configuration, punctuated by occasional transitions between such configurations; these transition events are infrequent in the sense that the average time between events is many vibrational periods. If a bottleneck region (saddle point) through which the system must pass in order to make the transition can be identified, the so-called transition state, then transition state theory (TST) can be used to separate the time scales between vibrations and transitions and statistically calculate the average amount of time the system would spend in a given state. In the transition state theory approximation, the classical rate constant for escape from the initial to the final state is taken to be the equilibrium flux through the dividing surface between initial and final state. Since atoms in crystals are usually tightly packed and the relevant temperatures are low compared with the melting temperature, the harmonic approximation to transition state theory can typically be used in studies of diffusion and reactions in crystals, and the problem reduces to determining the activation energy E_a and prefactor ν_0 of an Arrhenius-type expression for the rate k^{hTST} ,

$$k^{\text{hTST}} = \nu_0 \exp\left(-\frac{E_a}{k_B T}\right) \quad \text{with} \quad \nu_0 = \frac{\prod_{i=1}^{3N} \nu_i^{\text{valley}}}{\prod_{i=1}^{3N-1} \nu_i^{\text{saddle}}} \quad \text{and} \quad E_a = E^{\text{valley}} - E^{\text{saddle}}. \quad [1]$$

where E_a is the activation energy and ν_0 is the jump frequency. The jump frequency is calculated as the ratio of the product of the Γ -point frequencies of the ground state supercell over the product of all real Γ -point frequencies of the saddle point cell. Once all rates are known for diffusion and reaction of the relevant species in a system, the fast (vibrational) time scale can be discarded, and the long-time evolution of the system can be efficiently simulated with continuum (2) (“TCAD”) or kinetic-Monte Carlo (3) methods. Thus, one of the most important tasks for *ab-initio* calculations currently is to determine these kinetic parameters.

The calculation of the kinetic parameters starts with the search for the relevant saddle points and is one of the most challenging tasks in atomic-level process modeling. Only during the last decade, systematic methods based on transition state theory have been developed for a more reliable identification of saddle points. The currently most popular such method is the nudged elastic band method (NEB) (4), where a string of n images of the system is created and connected with springs in such a way as to form a discrete representation of a path from the initial to the final configuration. Initially, the position of each atom i in image j may be generated by a linear interpolation,

$$\mathbf{r}_i^j = \frac{n-j}{n} \mathbf{r}_i^{\text{initial}} + \frac{j}{n} \mathbf{r}_i^{\text{final}}. \quad [2]$$

An optimization algorithm is then applied to relax the images down towards the minimum energy path (MEP). The most straightforward approach would be to construct an object function

$$E_{\text{object}} = \sum_{j=1}^n E(\mathbf{r}_1^j, \mathbf{r}_2^j, \dots) + \sum_{j=1}^{n+1} \frac{k}{2} \left[\begin{pmatrix} \mathbf{r}_1^j \\ \mathbf{r}_2^j \\ \vdots \end{pmatrix} - \begin{pmatrix} \mathbf{r}_1^{j-1} \\ \mathbf{r}_2^{j-1} \\ \vdots \end{pmatrix} \right]^2 \quad [3]$$

($j = 0$ and $j = n + 1$ denote the initial and final states, respectively) and minimize with respect to the atomic positions in the intermediate images. This mimics an elastic band made up of n beads and $(n + 1)$ springs with spring constant k between the fixed initial and final configurations.

The elastic band tends to cut corners and gets pulled off the minimum energy path by the spring forces in regions where the minimum energy path is curved. Also, the images tend to slide down towards the endpoints, giving lowest resolution in the region of the saddle point, where it is most needed. These problems can be solved easily with a force projection. This is what is referred to as ‘nudging’. The reason for corner-cutting is the component of the spring force perpendicular to the path, while the reason for the down-sliding is the parallel component of the true force coming from the interaction between atoms in the system. Given an estimate of the unit tangent to the path at each image, the force on each image should only contain the parallel component of the spring force, and perpendicular component of the true force (4).

When multiple minimum energy paths are present between initial and final state, the optimization leads to convergence to the minimum energy path closest to the initial guess. In order to find the optimal minimum energy path in such a situation, some sampling of the various minimum energy paths needs to be carried out. The NEB method has been applied successfully to a wide range of atomic-scale process simulation problems, especially its implementation in the plane-wave density-functional theory code VASP (5). Examples are discussed in the following section.

Ab-Initio Calculation of Kinetic Parameters for Self-Diffusion: Si vs. Ge

The arrival of the nudged elastic band method has revolutionized the field of *ab-initio* calculations of kinetic parameters for process modeling, since the previously nearly exclusively used drag methods put large error bars on the calculated results and the predicted kinetic mechanisms. A considerable body of work has been produced during the past few years based on the nudged elastic band method, including bulk diffusion of impurity atoms in silicon such as boron (6) (including deactivation by clustering processes (7)), nitrogen (8), carbon (9) or phosphorous (10) and diffusion processes at and near semiconductor surfaces (11).

While the calculated values for such dopants (especially for dopants that diffuse with the help of self-interstitials) in general are predicted well by theory (15), the situation is different for the point defects, I and V , especially since also the experimental situation is very confusing. While the values for self-diffusion from tracer experiments agree well among each other and find $D_{\text{Si}} = 332 \exp[-(4.73 \text{ eV})/k_B T] \text{ cm}^2/\text{s}$ (2), even the most recent measurements do not agree whether I or V dominates self-diffusion at elevated temperatures. For example, recent values for self-diffusion activation range from $D_I C_I = 914$

$\exp[-(4.84 \text{ eV})/k_B T]$ cm²/s (21) to $D_I C_I = 149 \exp[-(4.68 \text{ eV})/k_B T]$ cm²/s (16) for I and from $D_V C_V = 0.6 \exp[-(4.03 \text{ eV})/k_B T]$ cm²/s (21) to $D_I C_I = 636 \exp[-(4.86 \text{ eV})/k_B T]$ cm²/s (16) for V , respectively.

The differences between experiments and theory are also dramatic. Density functional theory (DFT) calculations of diffusion in Si have involved both direct dynamical simulations (32) and static calculations of the formation and migration energy (24,33) using either the local density approximation (LDA) or the PW91 (36) generalized gradient approximation (GGA) functional. While the activation energy for Si self-diffusion calculated with PW91 is at the lower end of the nearly 1 eV range of experimental values (13), the most recent and supposedly reliable measurements give values at the high end (16,39).

In Ref. (27), a theoretical study of I defect energies has been reported, where LDA and GGA values have been compared to fixed-node diffusion quantum Monte Carlo calculations (DMC), which are supposedly more accurate since they are expected to describe electron correlation to a much better degree: DMC solves the many-body Schrödinger equation using stochastic methods without invoking a one-particle approximation like LDA and GGA. DMC finds an activation energy for neutral I of 5.4 eV (27), which is 54% (35%) higher than the corresponding LDA (GGA) value and 11–15% higher than the experimental values. Consideration of charged states can only lower the calculated neutral values. All this challenges the predictive capability of DFT/PW91 calculations for diffusion in semiconductors.

In the following, we will compare two different methods for possible “fixes” of traditional density-functional theory (DFT), applied to the cases of Si and Ge self-diffusion. Since DFT calculations are much more widely used and computationally significantly less expensive than DMC calculations, it is desirable to investigate if such fixes could decrease the differences between theory and experiment for self-diffusion in a systematic way. This is our topic of investigation in the following.

Self-Diffusion in Silicon

Predicting and controlling diffusion in semiconductors is of great importance in furthering the development of semiconductor technology. As the dimensions of circuits shrink, an understanding of the atomic-scale mechanism of diffusion processes becomes crucial in order to accurately model and design future devices. A prerequisite to understanding dopant diffusion is an understanding of defect mediated self-diffusion in semiconductors, since dopants – for example, Sb, Ge and As in crystalline Si – are believed to diffuse to a large extent with the help of defects, such as vacancies, in the crystal lattice (12,13). Accurate estimates of the activation energy and entropy of these defects are, therefore, essential.

In Ref. (14), Harrison has proposed a set of corrections to DFT calculations and applied them to the neutral tetrahedral self-interstitial (T) in Si. Since T has been found to be the transition state for neutral interstitial migration, the corrections that Harrison has derived apply directly to the activation energy of neutral I diffusion. The T interstitial ejects two electrons into the conduction band. Therefore, Harrison suggests a band gap correction for the too-small LDA band gap of 2×0.64 eV and an energy gain of -0.12 eV due to the formation of donor states for the two electrons. Due to the finite size of the 32-atom cell he considers, a part of the conduction band and the resonant p states in the conduction band are occupied, whereas in reality, the two electrons would reside in the donor states below the bottom of the conduction band. A finite-size correction gives rise to a -0.44 eV downward shift. In total, Harrison proposes a correction of $+0.72$ eV for the en-

ergy of the T interstitial. In Harrison's paper, no final value for the formation energy for the T interstitial is given.

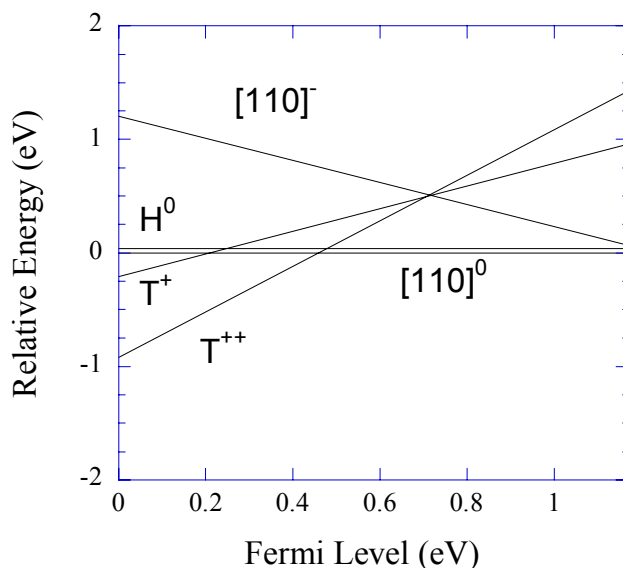


Figure 1. GGA-energetics of the silicon self-interstitial as a function of the Fermi level.

We apply analogous corrections (15) to 216-atom supercell calculations containing various point defect configurations in Si. In Fig. 1, we show the Fermi level dependence of the formation energy of an isolated Si self-interstitial in a 216-atom simulation cell. At low Fermi energies, the Si self-interstitial favors the double-positive charge state in the tetrahedral configuration (T^{++}). For $E_F > 0.45$ eV, the neutral [110] dumbbell split (X) configuration is dominant. The neutral H configuration is more or less degenerate with the X configuration, being only 0.04 eV higher in total energy which is well within our computational error bar. At diffusion temperatures, the intrinsic carrier concentration is typically high, and the Fermi level is in the n-type regime. The X interstitial thus can be assumed to represent the theoretical lowest-energy state of interstitials. Since the process of a diffusion hop itself is typically fast whereas changes in charge state depend on electron diffusion, we also assume that changes of the charge state of a defect during the diffusion process itself can be neglected, i.e., that the charge of the saddle point is identical to the charge of the initial (and thus final) state.

As concerning interstitial diffusion, we first discuss our results for the T interstitial. Within GGA, we find for the uncorrected formation energy a value of 4.01 eV. For the gap-correction with two electrons in the conduction band, we find 1.07 eV. For the finite-size correction, we find a finite-size correction of -0.35 eV. Band structure and density of states (DOS) [Fig. 2(a) and 3(a) and (b), respectively] seem to indicate that this cell is large enough to allow for the formation of the donor states, therefore we do not apply the respective (small) correction. We find a total correction of the T formation energy of +0.72 eV, which, added to the uncorrected formation energy of 4.01 eV, gives a corrected formation energy of 4.73 eV. This energy is 1.03 eV higher than the lowest-energy dumbbell (X) configuration, which thus is the value of the corrected migration energy, since X has no electrons beyond the valence band maximum [Fig. 2(b)]. The neutral activation energy of 4.73 eV is in excellent agreement with one of the recent measurements of the I activation energy of 4.68 eV (16).

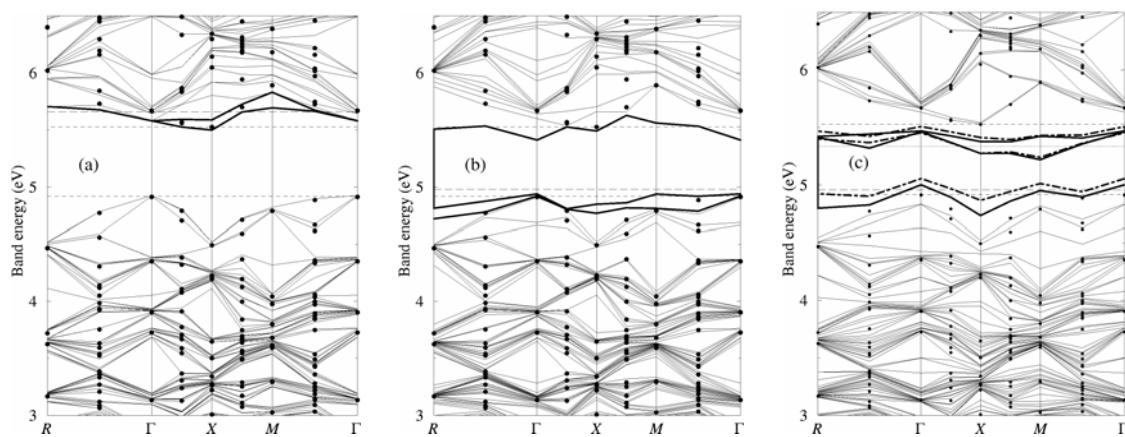


Figure 2. Band structure of 216-atom Si cells with a neutral (a) T interstitial, (b) X interstitial, and (c) vacancy. All structures are fully relaxed. Thin solid lines are the defective band structures, points the band structure of perfect Si-216. Long (short) dashed lines are Fermi levels (valance and conduction band edges of perfect Si). Gap bands of neutral (-1 charged) defects are shown as solid (dot dashed) fat lines.

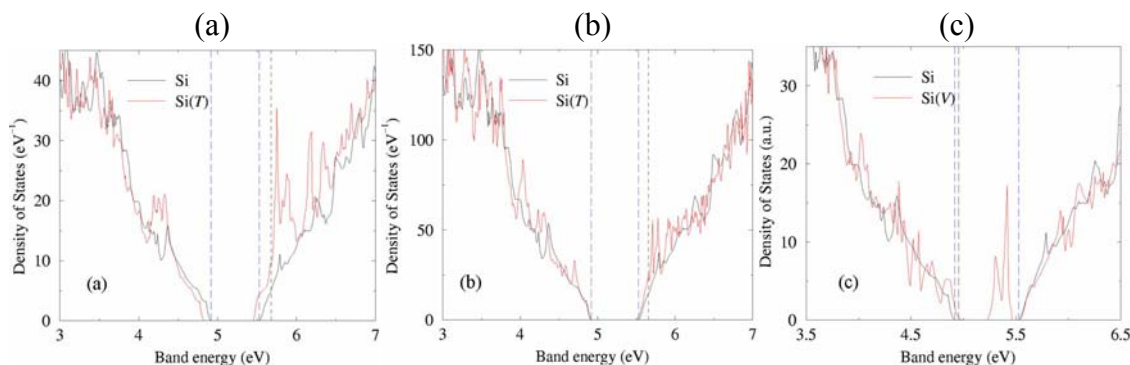


Figure 3. Electronic density of states for a tetrahedral interstitial in comparison to bulk Si in a (a) 64 and (b) 216-atom cell and a vacancy in a 216-atom cell (c). The long dashed lines indicate the band gap of bulk Si, the other dashed line indicates the (minimum) Fermi energy.

In the case of the vacancy, the technologically relevant charge states are currently under investigation (38), but a conclusive answer is not yet available. Thus, we examine in the following just the neutral charge state, which dominates most of the p-type and mid-gap regions [Fig. 4]. For the neutral unrelaxed V , simple perturbation theory (17) predicts three degenerate bands exactly at mid-gap which are occupied by two electrons. However, the Jahn-Teller (JT) theorem (18) predicts such a system to experience energy and symmetry lowering structure changes. Indeed, in neutral 216, 512, and 1000 atom cells with sufficient k -point sampling, we find next to no electrons above the Si valence band maximum, in agreement with previous estimates (19) [Fig. 2(c), solid gap bands in comparison to the interstitial band structures and Fig. 5 for the different cell sizes]. However, we want to point out that the introduction of charges changes the band structure as can be seen from a vacancy with negative charge [Fig. 2(c), dot dashed gap bands]. If the predicted JT shift were right and no levels above the valence band edge were occupied, the formation energy of the neutral V system (3.60 eV in GGA for sufficiently large cells) would not require any corrections. If we try to overcome the even larger lack of knowl-

edge about the gap states for the migration saddle by using Watkins' experimental value for neutral V migration of 0.45 eV (20), we end up with an upper-limit (GGA) activation energy of 4.05 eV. This compares favorably with the previous experimental value of 4.03 eV (21), but is considerably lower than the most recent measurement of 4.86 eV (16).

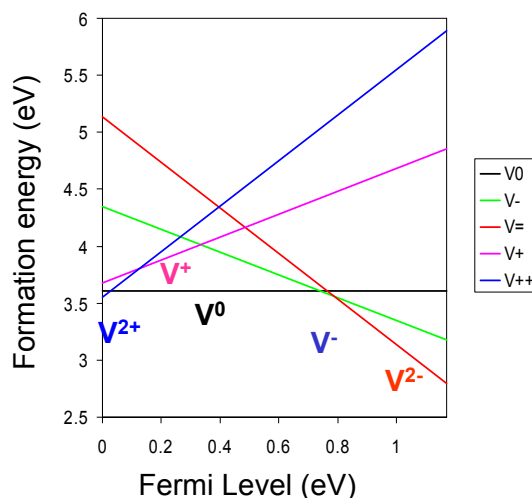


Figure 4. GGA-energetics of the silicon vacancy as a function of the Fermi level.

As a final note, we would like to point out that a 64-atom cell [Fig. 5(a)] produces defect states with very strong dispersion and usually does not produce Jahn-Teller distorted geometries, since a part of the bands is already below the valence band edge. Thus, no energy can be gained by distorting the symmetry. Therefore, neither the electronic properties nor volumetric relaxations, which are, e.g., needed to calculate the pressure dependence of vacancy formation, can be predicted correctly with such a cell (22).

In summary, although one could argue that the “ad-hoc” type fixes discussed in this section might be promising, the current results cannot give a positive evidence for their validity or non-validity.

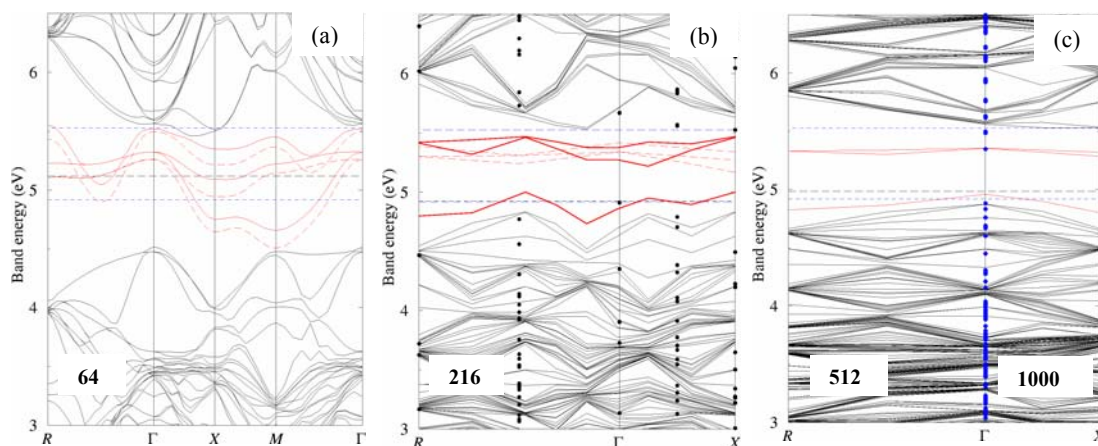


Figure 5. Band structure of a vacancy in a Si supercell with (a) 64, (b) 216, and (c) 512 (lines) and 1000 (dots at the Γ -point) atoms.

Self-Diffusion in Germanium

As Ge becomes more important as a material used in modern semiconductor devices, the efforts of better understanding its self-diffusion behavior are increasing. Various experiments, involving both ^{71}Ge tracers (13) and $^{70}\text{Ge}/^{74}\text{Ge}$ heterostructures (28), agree on an activation energy for self-diffusion (E_a) in the range of 2.95 to 3.14 eV and a large prefactor (D_0) of 8 to 44 cm^2/sec .

Consistent with early speculations (13), the observed diffusion behavior of Cu in Ge has shown convincingly that self-diffusion in Ge is vacancy mediated (29). This means that diffusion is a two step process, consisting of the creation and subsequent migration of a vacancy. Assuming both processes are thermally activated, E_a for diffusion is the sum of the formation and migration energies. Two different explanations have been given for the large magnitude of the prefactor: (i) that the vacancy is spread out over several atomic volumes resulting in a large entropy (30) and (ii) that force constants of the nearest neighbors are softened (31).

We have carried out DFT calculations for Ge self-diffusion using a periodic cell with 64 atoms ± 1 for the defects (23). Optimal atomic coordinates for stable configurations and saddle points for transitions between those stable configurations were obtained using the VASP code (5). At this cell size, the formation energy of the defects is converged to within 3% compared with 1000 atom cells (37). The PW91 functional was used in these calculations combined with ultrasoft pseudopotentials (40) and a plane wave basis set with an energy cutoff of 174 eV. A $4 \times 4 \times 4$ k-point grid (41) was used in the calculation of formation and migration energies, while harmonic frequencies were calculated with a $2 \times 2 \times 2$ grid.

The calculations presented here were done on charge neutral systems. DFT calculations are known to severely underestimate the band gap in most systems. The error in the calculated band gap for both Si and Ge is about the same amount (too low by 0.6 eV for Si, 0.8 eV for Ge). This error results in our DFT prediction that Ge is metallic (42). Thus, an explicit account of charge effects, known to be important in Si (43), is not possible. Because of this metallic nature, the simulated defects in Ge will effectively have the lowest energy charge state. This makes the explicit consideration of charge states unnecessary for predicting the lowest energy structures of the defects at this level of theory.

The cohesive energy of a Ge crystal is predicted quite accurately by the DFT/PW91 approximation. With respect to a free atom in the triplet state, the cohesive energy is predicted to be 3.73 eV, which agrees very well with the experimental value of 3.87 eV.

As described above, within harmonic TST, the rate of a process k depends on the activation energy E_a of the process and the harmonic prefactor ν (Eq. [1]). In turn, the diffusion constant for a given diffusion mechanism is described by

$$D = D_0 \exp\left(-\frac{E_a}{k_B T}\right), \quad [4]$$

where

$$D_0 = \frac{fz\nu_0^2}{2d} \exp\left(\frac{S_{\text{form}}}{k_B}\right). \quad [5]$$

In this equation, f is the correlation factor [24], z is the number of diffusion paths per site, a is the nearest neighbor separation, or the distance atoms move during a diffusion event,

and d is the dimensionality of the system. In the work described below, $z = 4$, $a = 2.5 \text{ \AA}$, and $d = 3$. The value of f depends on the specific diffusion mechanism.

We have used the climbing-image nudged elastic band (CI-NEB) method for finding the MEP (45). This is a modification of the NEB where one of the images is made to converge rigorously to the highest saddle point. The diffusion prefactor D_0 depends on the entropy of formation (S_{form}) and migration (expressed here as the rate prefactor ν_0). These were calculated by constructing the force constant matrix for the relevant geometries (the perfect crystal and the minimized defect structure for S_{form} and the defect structure and the saddle point for ν_0) (46). The second derivatives of energy with respect to atomic coordinates were evaluated using finite differences of 0.01 \AA . Only atoms within a specified radius of the diffusing atom(s) were displaced. The resulting matrix was diagonalized to find the normal mode frequencies needed for the rate calculation. The calculation was repeated for different values of the radius to ensure convergence.

In the following, we only discuss the theoretically and experimentally dominant vacancy mediated diffusion in detail, and only quote the results for interstitial-mediated diffusion and direct exchange in Table II (23).

The formation energy is predicted by PW91 to be 2.22 eV . The migration path is simple: one of the four neighbors moves into the vacant site. The migration barrier is low, 0.19 eV , giving an overall E_a of 2.41 eV . We find the structural relaxation around the vacancy to be quite local: Whereas the neighboring four atoms relax inward by 0.36 \AA toward the vacant lattice site, the second neighbors only move by 0.02 \AA . The same is not true for the force constants, which are affected over many neighbor shells. Table I shows how the formation entropy converges with the size of the region included in the calculation. An infinite cutoff extrapolation is obtained, assuming an exponential dependence (47). The vibrational entropy of formation is calculated to be large, $S_{\text{form}} = 9.18 k_B$. The main reason for this large entropy is the softening of the force constants, similar to what was assumed in Ref. (31) except that a large region around the vacancy is affected, not just the adjacent bonds.

TABLE I. Vibrational entropy of defect formation calculated with various values of the cutoff radius. Extrapolated, infinite size limit is obtained by fitting an exponential.

Vacancy		Interstitial	
Number of atoms	Entropy (k_B)	Number of atoms	Entropy (k_B)
4	3.54	2	1.56
16	6.43	6	2.43
28	7.84	18	3.37
		30	3.66
Extrapolated	9.18	Extrapolated	3.70

The migration prefactor is calculated to have a magnitude of $\nu_0 = 2.40 \times 10^{12} \text{ s}^{-1}$, which is typical for prefactors. The total diffusion prefactor is then $D_0 = 4.85 \text{ cm}^2 \text{ s}^{-1}$, where $f = 1/2$. This result is close to but slightly smaller than the range of experimental estimates, see Table II. The overall activation energy for diffusion (formation energy of the vacancy plus migration energy), is however predicted to be $\sim 0.7 \text{ eV}$ too small.

The PW91 results predict that vacancy-mediated diffusion, the mechanism determined experimentally to dominate self-diffusion in Ge, is indeed faster than the other two mechanisms studied, but the overall rate of diffusion is predicted to be too fast compared to experiment as can be seen in Fig. 6. The disagreement is due to the activation energy, either the formation energy of the vacancy or the migration energy or both. The theoretical prefactor agrees well with the experimental one. Provided that the experimental numbers are not tainted by events such as V - I recombination and thus give the “true” V diffu-

sivity, the most likely source of error in the activation energy is the approximate nature of the exchange-correlation functional used in the DFT calculations. In the following, we estimate what the B3LYP functional (34) would give for the activation energy using clusters carved out of the periodic system (23). We find a B3LYP correction of just 0.07 eV to the migration energy of the vacancy mechanism, but a large 0.56 eV correction to the formation energy, resulting in a corrected activation energy of $2.41 \text{ eV} + 0.63 \text{ eV} = 3.04 \text{ eV}$, which agrees very well with experiment. However, this correction procedure predicts a cohesive energy for Ge that is in worse agreement with experiment than the PW91 results, lowering the PW91 value by 0.63 eV, from 3.73 eV to 3.10 eV. Thus, also this kind of “DFT”-repair leaves questions unanswered, leaving self-diffusion in semiconductor a worthy field of theoretical investigations even for the future.

TABLE II. Calculated values for the diffusion coefficient and activation energy for the vacancy (V), interstitial (I) and concerted exchange (CE) in Ge, as well as experimental values (23).

Quantity	V	I	CE	Experiment
$D_0 \text{ (cm}^2\text{/s)}$	4.85	0.08	1.13	8 to 44
$E_a \text{ (eV)}$	2.41	3.77	3.56	2.95 to 3.14

Conclusions

In this paper, we have shown concepts and examples of ab-initio based methods to further our understanding especially of kinetic parameters, which can be determined with increasing accuracy using modern methods such as the nudged elastic band method. We have also demonstrated the possible errors in such calculations and have examined possible fixes to the current theory. The accuracy of the calculations remains unclear, leaving more final answers to the questions about self-diffusion in semiconductors to the expected advent of more sophisticated exchange-correlation functionals in main-stream ab-initio codes in the near future.

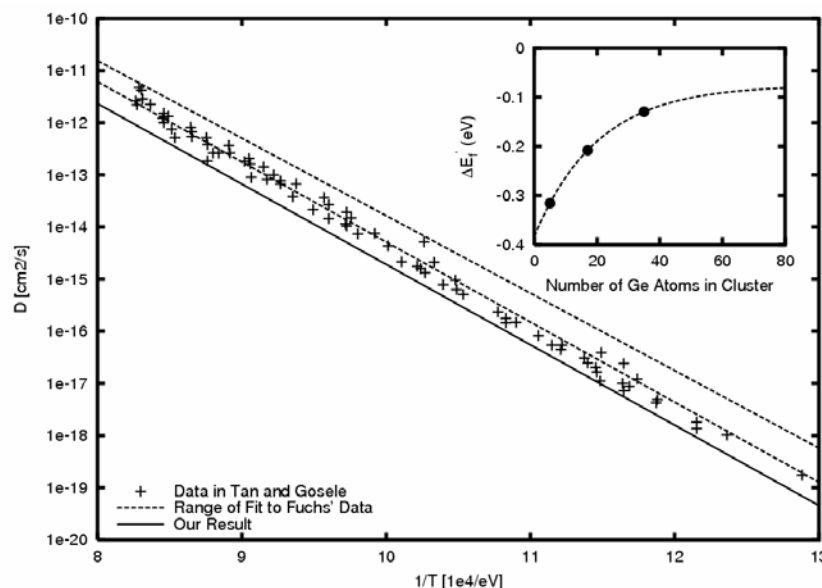


Figure 6. Comparison between the calculated results (solid line) and experimental data: The two dashed lines give a maximum and minimum estimate from Ref. (13), pluses (+) give values from Ref. (25). Inset: Estimated B3LYP correction to the vacancy formation energy as a function of cluster size. The dashed line is an exponential fit to the data, giving an infinite-cluster limit of -0.07 eV .

References

1. W. Windl, Phys. Stat. Sol. (B) 241, 2313 (2004).
2. P. Pichler, *Intrinsic Point Defects, Impurities, and Their Diffusion in Silicon* (Springer, 2004), Chapter 1.5.
3. N. Strecker, V. Moroz, and M. Jaraiz, Proceedings of the 2002 International Conference on Computational Nanoscience, p. 247 (2002).
4. H. Jónsson, G. Mills, and K. W. Jacobsen, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne *et al.* (World Scientific, Singapore, 1998), p. 385.
5. G. Kresse and J. Hafner, Phys. Rev. B 47, 558 (1993); 49, 14 251 (1994); G. Kresse and J. Furthmüller, Comput. Mater. Sci. 6, 15 (1996); Phys. Rev. B 54, 11 169 (1996).
6. W. Windl, M. M. Bunea, R. Stumpf, S. T. Dunham, and M. P. Masquelier, Phys. Rev. Lett. 83, 4345 (1999).
7. X.-Y. Liu, W. Windl, and M. P. Masquelier, Appl. Phys. Lett. 77, 2018 (2000).
8. N. G. Stoddard, P. Pichler, G. Duscher and W. Windl, Phys. Rev. Lett. 95, 025901 (2005).
9. C. L. Liu, W. Windl, L. Borucki, S. F. Lu and X. Y. Liu, Appl. Phys. Lett. 80, 52 (2002).
10. X.Y. Liu, W. Windl, K.M. Beardmore and M.P. Masquelier, Appl. Phys. Lett. 82, 1839 (2003).
11. B. P. Uberuaga, M. Levskovar, A. P. Smith, H. Jónsson, and M. Olmstead, Phys. Rev. Lett. 84, 2441 (2000).
12. J. D. Plummer, M. D. Deal and P. B. Griffin, *Silicon, VLSI Technology* (Prentice Hall, 2000).
13. W. Frank, U. Gösele, H. Mehrer, and A. Seeger, in *Diffusion in Crystalline Solids*, edited by G. E. Murch and A. S. Nowick (Academic Press, Orlando, 1984), p. 63.
14. W. A. Harrison, in *Defects and Diffusion in Silicon Processing*, ed. by T. Diaz de la Rubia *et al.*, MRS Proc. 469 (MRS, Pittsburgh, 1997), p. 211.
15. W. Windl, Phys. Status Solidi B 241, 2313 (2004).
16. A. Ural, P.B. Griffin, and J. D. Plummer, Phys. Rev. Lett. 83, 3454 (1999).
17. H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).
18. H. A. Jahn and E. Teller, Proc. Roy. Soc. A 161, 220 (1937).
19. G. A. Baraff, E. O. Kane, and M. Schlüter, Phys. Rev. B 21, 3563 (1980).
20. G. D. Watkins, Phys. Solid State 41, 746 (1999).
21. U. Gösele, A. Plössl, and Y. T. Tan, in *Process Physics and Modeling in Semiconductor Technology*, ed. by G. R. Srinivasan, C. S. Murthy, and S.T. Dunham (Electrochemical Society, Pennington, NJ, 1996), p. 309.
22. W. Windl, M. S. Daw, N. N. Carlson, and M. Laudon, in *Advances in Materials Theory and Modeling - Bridg-ing over Multiple Length and Time Scales*, ed. by V. Bulatov *et al.*, MRS Proc. 677 (MRS, Pittsburgh, 2001), p. AA9.4.1-6.
23. B. P. Uberuaga, G. Henkelman, H. Jónsson, S. T. Dunham, W. Windl, and R. Stumpf Phys. Stat. Sol. B 233, 24 (2002).
24. R. J. Needs, J. Phys.: Cond. Matter 11, 10437 (1999).

25. T. Y. Tan and U. Gösele, in *Handbook of Semiconductor Technology-Electronic Structure and Properties of Semiconductors*, ed. K. A. Jackson and W. Schroeter, vol. 1 (Wiley & Sons, New York, 2000), p. 231.
26. M. Puska, S. Pöykkö, M. Pesola, and R. Nieminen, *Phys. Rev. B* 58, 1318 (1998).
27. W.-K. Leung, R. J. Needs, G. Rajagopal, S. Itoh, and S. Ihara, *Phys. Rev. Lett.* 83, 2351 (1999).
28. H. D. Fuchs, W. Walukiewicz, E. E. Haller, W. Dondl, R. Schorer, G. Abstreiter, A. I. Rudnev, A. V. Tikhomirov, V. I. Ozhogin *Phys. Rev. B* 51, 16817 (1995).
29. N. A. Stolwijk, W. Frank, J. Hölzl, S. J. Pearton, and E. E. Haller, *J. Appl. Phys.* 57, 5211 (1985).
30. A. Seeger and K. P. Chik, *Phys. Status Solidi* 29, 455 (1968).
31. J. C. Burgoin and M. Lanoo, *Radiat. Eff.* 46, 157 (1980).
32. P. E. Blöchl, E. Smargiassi, R. Car, D. B. Laks, W. Andreoni, and S. T. Pantelides, *Phys. Rev. Lett.* 70, 2435 (1993).
33. W. C. Lee, S. G. Lee, and K. J. Chang, *J. Phys.-Condensed Matter* 10, 995 (1998).
34. A. D. Becke, *Phys. Rev. A* 38, 3098 (1988), C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B* 37, 785 (1988).
35. B. P. Uberuaga, Ph. D. thesis, University of Washington (2000). The PW91 predicted activation energy for diffusion in Si is 4.57 eV for CE, 4.04 eV for I, and 3.83 eV for V. The cohesive energy of Si is 4.57 eV with PW91, compared with an experimental value of 4.62 eV.
36. J. P. Perdew, in *Electronic Structure of Solids*, eds. P. Ziesche and H. Eschrig (1991).
37. W. Windl (unpublished).
38. H. Bracht, *Physica B: Physics of Condensed Matter* 376, 11 (2006).
39. H. Bracht et al., *Phys. Rev. Lett.* 81, 393 (1998).
40. D. Vanderbilt, *Physical Review B* 41, 7892 (1990); [20] G. Kresse and J. Hafner, *J. Phys: Cond. Matter* 6, 8245 (1994).
41. H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* 13, 5188 (1976).
42. A. Seidl, A. Gorling, P. Vogl, J. A. Majewski, and M. Levy, *Phys. Rev. B* 53, 3764 (1996).
43. W. A. Harrison, *Phys. Rev. B* 57, 9727 (1997).
44. A. D. LeClaire in *Diffusion Processes*, ed. J. N. Sherwood, et al. (Gordon and Breach, London, 1971).
45. G. Henkelman, B. P. Uberuaga and H. Jónsson, *J. Chem. Phys.* 113, 9901 (2000); 113, 9978 (2000).
46. G. H. Vineyard, *J. Phys. Chem. Solids* 3, 121 (1957).
47. We use an exponential function here just for convenience. The cluster size dependence must converge to some limiting value and an exponential is the simplest function with such a shape.
48. K. C. Pandey, *Phys. Rev. Lett.* 57, 2287 (1986).
49. We note that previous LDA calculations including only the Γ -point gave even lower defect formation energy than PW91, see A. Janotti et al., *Physica B* 273-274, 575 (1999).